B·O·S·C

BOARD OF SCIENTIFIC COUNSELORS

December 29, 2009

Dr. Paul Anastas
Assistant Administrator
Office of Research and Development
U.S. Environmental Protection Agency

Dr. Robert Kavlock
Director
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Dear Dr. Anastas and Dr. Kavlock:

This is a letter report from the Board of Scientific Counselors' (BOSC) review of the Office of Research and Development's (ORD) Computational Toxicology Research Program (CTRP). The BOSC Computational Toxicology Review ~~Subcommittee~~Committee reviewed CTRP's progress and future plans during a conference call on September 25, 2009, and a 2-day meeting held September 29-30, 2009, in Research Triangle Park, North Carolina. The members of the BOSC Review ~~Subcommittee~~Committee are George Daston (Chair), James Clark, Richard DiGiulio, Ali Faqi, Lawrence Hunter, Moiz Mumtaz, Dennis Paustenbach, John Quackenbush, Santiago Schnell, Cynthia Stokes, and Katrina Waters.

This is the fourth review of the CTRP conducted by the BOSC. The National Center for Computational Toxicology (NCCT) first became operational in February 2005. During the 4.5 years between its establishment and this review, the CTRP has made substantial progress in establishing and meeting its priorities and goals, collaborating within and outside EPA to leverage the staff's expertise and transforming the field of toxicity testing. Many of the recommendations made by the BOSC during its earlier reviews have been acted on by the CTRP. This includes improved capabilities in bioinformatics through the funding of two external centers and in informatics and systems biology through staff hires; expansion of the CTRP's technical approaches to even more programs within the Agency; and the formation of an extensive collaboration with the National Institute of Environmental Health Sciences (NIEHS) and the National Human Genome Research Institute (NHGRI) for its ToxCast project.

Commented [J1]: Remove line numbering

Commented [J2]: Should this always be referred to as the Review Committee to avoid confusion? Or can it bedefined at first use as "the committee"

Does it need to be caps?

I thought it was agreed that we should not call these subcommittees, but rather just review committees

The purpose of the September 2009 review was to provide the CTRP with advice on (1) the progress the Center has made, in the past 4.5 years, in fulfilling its mission and strategic goals; and (2) whether the NCCT should continue as an established organization beyond its original 5-year charter. In particular, the BOSC addressed five charge questions that focused on the progress and future of the NCCT. The BOSC's responses to these questions follow.

Charge Question 1: *What is your evaluation of the progress the CTRP has made in achieving its original goals and objectives, and whether it has efficiently utilized available resources?*

The mission statement of the CTRP is to integrate modern computing and information technologies with molecular biology to provide the Agency with decision support tools for high-throughput risk assessment.

The three initial long-term goals of the CTRP were as follows:

1. Risk assessors use improved methods and tools to better understand and describe the linkages of the source-to-outcome paradigm,

2. EPA Program Offices use advanced hazard characterization tools to prioritize and screen chemicals for toxicological evaluation, and

3. EPA assessors and regulators use new and improved methods and models based on the latest science for enhanced dose-response assessment and quantitative risk assessment.

These are ambitious goals for the Program, particularly the last one, the objective of which is to identify non-classical methods for assigning an approximate acute and chronic toxicological hazard rating to hundreds, if not thousands, of chemicals. The vast majority of these ratings are to be conducted without the traditional reliance on animal testing and human epidemiology.

The BOSC members believe that the CTRP has made substantial progress toward meeting the original long-term goals, and that the progress is appropriate given the duration of the Program's existence and the resources involved. Although the goals have not been fully met, the CTRP has made significant advancements to creating and providing the methods, tools, and models that will enable risk assessors, program offices, and regulators to use 21st century science and technology for their work, as indicated in the goals. Notably, a number of new tools already have been released to support decision-making, including AcToR, DSSTox, ToxRefDB, and ExpoCast.

Underpinning all of these, the CTRP has focused on integrating modern computational approaches with molecular and cellular biology and physiology to create new methods for EPA's chemical prioritization and risk assessment efforts. The activities of the CTRP in pursuit of these goals have included the assembly and integration of vast quantities of existing toxicological and toxicogenomics data; creation of new database and data warehousing tools to house the data; development of methods and tools for discerning actionable knowledge from the data; additional data acquisition on chemicals of Agency interest; and the development of various types of computational models to understand biological and toxicological mechanisms and provide

predictive tools for hazard evaluation and prioritization and risk assessment. As several of the CTRP's projects have matured, they have started to link together in ways that allow ToxCast and Tox21 data to be effectively captured and managed, and allow the various sources of information to be leveraged against each other in productive ways.

Overall, during its inaugural funding period, the CTRP has built the infrastructure necessary to bring computational tools to risk assessment; assembling the data and building the tools that are needed to collect the high-throughput screening (HTS) data that the Program is now facing. In the process, the Program staff has learned the limitations of the existing data and has begun to explore how it might address the limitations of the existing systems, data, and models that will be necessary to move forward. Issues of bioavailability and bioactivity, correlations between *in vitro* and *in vivo* results, and dose response relationship and toxicity are cases in point.

The BOSC concludes that the CTRP has laid a strong foundation and has put forth good ideas to move forward, and that there is clearly a need to continue to build on the success of the Program.

Because of the large number of projects and collaborations involved, it is difficult to assess in detail whether the CTRP has utilized the available resources efficiently. Nonetheless, it is the conclusion of the BOSC that the progress made is substantial and appropriate in light of the time the Program has been in existence, its budget, the number of personnel involved, and the collaborations in place within and outside the Agency.

As expected with such an ambitious program, the CTRP has been faced with a number of challenges as the Program has strived to meet its goals. The following are the BOSC's observations about some of these challenges along with suggestions for meeting them that are relevant to continued progress.

One of the challenges that the CTRP has taken on is the assembly and integration of the vast quantities of existing, available toxicological and toxicogenomics data. Although it might seem that this should be a relatively easy task, much of the data are not electronic, much is poorly annotated, and many ambiguities exist within the available data. The CTRP has initiated a number of data base and data warehousing projects to begin to address these issues, including AcToR, DSSTox, ToxRefDB, and ExpoCast, each of which approaches a different aspect of the problem. Each of these has made significant progress over the past few years, along the way discovering the need to do a significant amount of manual curation, but using existing tools and developing new ones as appropriate, and they are beginning to be linked together in productive ways. Despite all that has been achieved, much of the work is clearly in an early stage, and many of the current resources are more like a structured index to the underlying data rather than a comprehensive linked data resource. These projects need to continue to build on things that are in place, drilling deeper into the data and continuing the problem of structuring, standardizing, and organizing the data so that they can be more easily subjected to comprehensive meta-analyses. In addition, much of what is available is focused on research scientists and the data are presented "as is" without context. The potential problem is that as these data resources are publicly available, the public may access the information and, without context, may misinterpret what is there. For example, the data in ACToR identifies compounds that have been tested for carcinogenicity or genotoxicity, but not those that have been found to be carcinogenic or

genotoxic. In this context, it might be worthwhile to obtain some public feedback on how people may interpret the available data.

~~Although~~ ~~D~~database development (sometimes called knowledge warehouse) ~~is seen as a fairly mundane task, it~~ is the foundation on which much of the CTRP is being built and doing it successfully requires a significant intellectual investment. Fundamentally, a database is a model, and building the model requires two components: understanding the relationship between the data elements and understanding how people will use the data. In many ways, the CTRP is in a critical phase in that, having assembled the data into one place, building the linkages in a systematic fashion now will require understanding the value of data to the community and the manner in which those data will be used.

Much~~A large part~~ of the success of the CTRP relies on the development of computational models to interpret data and make predictions. A major part of the modeling effort focuses on interrogating the databases. The ~~Subcommittee~~Committee noted that a substantial part of these efforts utilizes machine-learning methods. Although the BOSC has no specific objection to such approaches, it should be noted that many biostatisticians are sometimes apprehensive about such methods. Because many of the ultimate customer clients and stakeholders of the CTRP's efforts are expected to have biostatistical backgrounds, the BOSC encourages the Program to consult with biostatisticians early and often to assure they can address any objections; the CTRP also should consider attempting some additional methods.

There also is a need to interact more extensively with the broader scientific user community in the process of developing and rolling out tools and software. One of the driving principles behind the program is that the resources it produces should be both useful and used. For the resources to be useful, they must address important questions; the CTRP seems to have a strong focus on addressing relevant problems in environmental toxicology and exposure and risk assessment. For the resources to be used, the Program must understand how risk assessors outside the CTRP and within the broader toxicological and toxicogenomics community will use these resources. This process should allow the development of appropriate use cases that can guide how the tools are created and integrated. If this is done in a systematic manner, it has the potential to rapidly advance the evolution of the resources the CTRP is developing. This could be achieved through an annual or biannual conference by bringing together the data generators, the data users, and the risk assessors/managers—the ultimate users of these alternative methods/models.

There were concerns expressed by some ~~Subcommittee~~Committee members that associations are not causation and this should be recognized by the EPA management, both at the CTRP level and at the level of the Office of the Administrator. ~~-~~The results of a computer-generated association should be carefully examined through traditional testing and careful scientific study. ~~-~~Some of the members of the BOSC Review Committee recalled that the Ames test once was thought to serve in a similar manner as that being proposed for the various tools that are being built by the CTRP. ~~–~~Thus, although the BOSC fully supports providing more resources to CTRP efforts, it offers the precautionary warning that, at best, the results of these efforts will be the temporary placement of a chemical into a bin that could likely initiate a "science forcing event." ~~–~~In turn, the manufacturer or user of the chemical can be put on notice that this chemical appears to have certain characteristics that give it likelihood for being a hazard and that they will need to

conduct further toxicological testing. This will bring the chemical industry on par with the pharmaceutical industry that has been conducting such cost/risk/benefit analyses for several decades.

Recommendations:

- ✧ Several CTRP projects have undertaken structuring, standardizing, and organizing the data so that they can be more easily subjected to comprehensive meta-analyses. At this point, the CTRP should obtain some public feedback on how people are using and interpreting the available data.

- ✧ Acceptance of products, methods, and databases by the risk assessment community is the key to success. Hence, the NCCT should organize an annual or biannual conference that brings together the data generators, data users, and risk assessors/managers—the ultimate users of these alternative methods/models.

- ✧ As more data from high-throughput assays and computer models become available, the NCCT should provide guidance on how to interpret this information in the context of more traditional testing and scientific examination so that risk assessment practitioners in the EPA program offices can apply these findings.

Charge Question 2: *To what extent and how effectively has the CTRP utilized internal and external partnerships to foster its goals?*

It is clear that the CTRP staff has been effective at finding professional colleagues in various institutions with whom to collaborate. Certainly, the Program has been able to identify various research activities and data sets through EPA and other organizations in an attempt to assemble sufficient information to achieve some of its goals. For example, the exposure assessment work by Hubal, et al., is an example of a group that has a good idea about where various data sets reside in the government and they are committed to make more of those data sets available. When those data are gathered, they will be accessed by other groups in the CTRP to combine with data the Program has on toxicity or predictors of toxicity.

It appears that the CTRP has successfully engaged those in the NCCT, National Health and Environmental Effects Research Laboratory (NHEERL), National Exposure Research Laboratory (NERL), National Risk Management Research Laboratory (NRMRL), National Center for Environmental Assessment (NCEA), and other groups within the Federal Government. There are two major government agencies, however, which support research work relevant to the Computational Toxicology program— the National Science Foundation (NSF) and the Department of Energy (DOE). These agencies do not appear to be CTRP partners. NSF has basic science programs that are relevant not only to the biological sciences, but also computer science and the area of qualitative research into the activities (and information needs) of scientists. These are essential components of the CTRP research. DOE programs and data, particularly in radiation safety and environmental remediation, would seem natural sources of valuable information for the CTRP. In addition, the collaboration with the National Institutes of

Health (NIH) appears to be mediated by NIEHS and the NIH Chemical Genomics Center (NCGC). Also, a more formal relationship can be established with the National Library of Medicine, particularly with the National Center for Biotechnology Information and its PubChem program. The BOSC members believe that the development of these relationships would be appropriate for the CTRP.

It appears that the CTRP has been less effective in developing equally strong research groups at various universities in the United States and internationally, or partnerships with other scientific and regulatory bodies outside the United States (e.g., the United Kingdom, Australia, New Zealand, Switzerland, Germany, and a few other countries have substantial databases on exposure, toxicology, and predictive toxicology). No doubt, these relationships will follow as the CTRP matures and receives additional funding. At the moment, the Program is establishing collaborations in multi-scale modeling of developmental toxicity and virtual tissues with some U.S. academic partners (Indiana University and University of Texas). Although these collaborations clearly are of value, it also seems appropriate to expand these partnerships to include multi-scale modeling work in Europe, Asia, and elsewhere in the United States. An excellent example is the virtual physiological human initiative that is intended to support the development of patient-specific computer models and their applications in personalized and predictive medicine. This constitutes an integral part of the international Physiome Project, a worldwide public domain effort to develop a computational framework for quantitative description of biological processes in living systems across all relevant levels of structural and functional integration, from molecule to organism, including the human. The Physiome Project has established standards, which are used in the multi-scale modeling community. In parallel, the mathematical and computational oncology community has established similar initiatives for assessing cancer progression and treatments. The CTRP will benefit substantially in establishing partnerships with these multi-scale modeling enterprises in the United States, Europe, and Asia.

As an advisory group, the BOSC would very much like to have seen a table which presented the various relationships with the full-time equivalents (FTEs) from each organization committed to a particular "joint" collaboration. It would have been helpful to include in that table an indication of the level of financial resources from the CTRP that were dedicated to the various projects, and a timeline for various milestones. Like so many government initiatives, it is very difficult to rate the efficiency or productivity of the Program because it is unclear if the project's success is the result of CTRP resources or those of another entity. For example, if the CTRP is devoting 0.25 FTE for 3 years to the initiative and NHEERL is devoting 5 FTEs from other programs for 4 years, the success of the project is predominantly the result of NHEERL resources. Another layer of complexity occurs when some portion of the project is conducted by post-docs, summer or more permanent interns, and graduate students (who may or may not be counted as FTEs). Then, beyond that, there are contractors who often participate in projects. Building such teams is not to be discouraged and, indeed, it should be promoted. It is not possible, however, to determine which person or persons is "leading the project" and providing the driving force to resolution. Only by interviewing each team, would it be clear who was truly providing the leadership and effort~~laboring ore~~.

Recommendations:

- ✧ Continue to interact with other scientific bodies, regulatory agencies, and universities both in the United States and globally so as to insure that work conducted elsewhere can be "built upon." In addition, it is recommended that the group interact with the toxicology groups within pharmaceutical and major chemical companies. One possible benefit of this interaction is that it may promote harmonization regarding the organization of historical data that currently are being assembled, as well as new data. This would eliminate the time-consuming task of extracting data from original studies and then entering them in the databases.

- ✧ Routinely (perhaps biannually) sponsor some sort of exchange of information with risk assessment practitioners both inside and outside EPA (corporations, consultants, and government scientists) to be sure that the end products of the Program's work are both reliable and of use to the future users.

- ✧ For the next BOSC review, develop a table that presents the level of effort dedicated to specific projects, by year. This table would contain the number of CTRP FTEs, as well as the approximate level of "collaborative" effort (from other EPA laboratories and other partners and consultants). In kind support and "hard" dollars also should be presented.

Charge Question 3: *What evaluation can you provide relative to the contributions of the CTRP to the advancement of transforming the field of toxicity testing?*

The CTRP appears to be at the cutting edge of transforming the field of toxicity testing. Its original goals, defined at its creation in 2005, were ~~highly~~ consistent with recommendations subsequently described in the National Academy of Sciences' 2007 report, entitled *Toxicology in the 21st Century: A Vision and a Strategy*. Particularly important contributions in this area include major advances in HTS, advances in approaches for data-mining from various data sources including HTS and other ToxCast efforts, advanced model development of virtual tissues, and the incorporation of uncertainty analysis into model development and, ultimately, risk assessments.

The incorporation of modern computing with molecular biology as developed by the CTRP is necessary to move the discipline of toxicology from the current stage, which is primarily descriptive science, to a more predictive one. The concept of utilization cell-based *in vitro* testing will assist the understanding of the key biological pathways by which chemicals induce adverse effects. The development of knowledge bases of toxicity pathways, toxicological responses, and key information on biological networks will lead to the use of solid science in the risk assessment. The HTS will lead to more cost-effective testing, which will save mon~~e~~y and reduce the use of animal testing. In addition, virtual tissues (liver and embryo) that link across levels of biological organization from molecular to cellular to tissue level responses will be good predictive tools for general and developmental toxicity. The NCCT is establishing collaborations with other institutions across the world with similar goals, which may lead to the expansion of the number of predictive virtual tissues. Moreover, it is expected that the attrition rate for pharmaceutical compounds will be reduced as the computational toxicology tools

provide a better prediction of human toxicity.

The BOSC applauds the efforts of the NCCT to embark on the challenges of the Tox21 paradigm, elucidating the strengths and limitations of the "toxicity pathway" approach and the challenges of generating truly predictive HTS platforms. The CTRP has demonstrated rapid progress in the development of HTS. To date, during Phase I of ToxCast, 467 assays have been employed with nine platforms. These assays have been applied to an initial set of 309 chemicals for which relatively extensive toxicological information is available (largely active chemicals in pesticides). In Phase II, this number of assays is expanding and will be applied to a new set of 700+ chemicals for which less information is available. This phase will provide for a critical evaluation of the large array of assays under consideration for their relative utility of chemical prioritization. The challenge for this group is to apply the lessons learned from the Phase I ToxCast efforts and iterate with Phase II to add new assays and define a strategy for attrition of those that provide limited or inconsistent information. The BOSC recommends that the Program keep the statisticians and mathematical modelers involved in assay evaluation so that they can move from qualitative prediction to quantitative prediction of outcomes from exposure data. Although a strategy has been outlined to include this information, it has not been sufficiently communicated to all members of the team. In addition, to achieve their future milestones related to prediction of outcome, the group needs to define metrics by which they can measure their success (such as specificity and sensitivity goals) and declare victory for specific classes of compounds or cell types. The identification of new toxicity pathways, and therefore new assays, will be essential to gain the predictive power necessary to predict outcome from exposure beyond a single class of compounds. Again, a strategy to define new pathways/assays may exist but, when questioned at the poster session, many of the Program researchers were unaware of who was doing it or how it was being done.

The initial evaluation of Endocrine Disruptor profiling with ToxCast data is extremely promising in its ability to identify new modes of action (MOA) that are not traditionally considered in addition to known developmental or reproductive endpoints. Likewise, the ToxCast data display good sensitivity for predicting neoplastic liver lesions, and statistical models have demonstrated a data gap for new assays to be developed for non-neoplastic lesions. Although there is some suspicion of these statistical approaches because "they don't make sense based on what we know", we recommend an unbiased evaluation of the usefulness of particular assays to achieve prediction beyond a single class of compounds and to define knowledge gaps for new assay design. This is a huge step forward for the NCCT to demonstrate the potential utility of the ToxCast approach, and it justifies continued funding to maintain the momentum of this research team.

The engagement of collaborators in NCEA to transform risk assessment into a "NexGen" paradigm using tools and databases coming out of the NCCT is certainly impressive, although daunting. Translating the predictions from the HTS assays to human population risk will require strong connections between the MOA data and statistical genetic diversity, such as those being provided by the Carolina Center for Computational Toxicology, led by Dr Ivan Rusyn. The ability to put real uncertainty factors into exposure limits that protect 99 percent of the population will truly revolutionize human health risk assessment.

It is very difficult to determine how the work of the CTRP will be implemented from a regulatory standpoint. The Program's work will likely help organizations identify the chemicals most deserving of significant study and the "type" of additional testing that needs to be conducted. Ultimately, the CTRP research will establish a methodology that can be relied upon by researchers around the world for quickly identifying those chemicals that "have a red flag." The BOSC recommends that the NCCT develop case studies that demonstrate a strategy for incorporation of CTRP tools/research into the risk assessment process. This could help significantly in optimizing the resources of the various organizations such that they will not be conducting routine toxicology "screening tests" on chemicals that raise too many red flags after being run through the ultimate program offered by the CTRP.

Recommendations:

- ✧ Keep the statisticians and mathematical modelers involved in assay evaluation so that they can move from qualitative prediction to quantitative prediction of outcomes from exposure data.

- ✧ Conduct an unbiased evaluation of the usefulness of particular assays to achieve prediction beyond a single class of compounds and to define knowledge gaps for new assay design.

- ✧ Develop case studies that demonstrate a strategy for incorporation of CTRP tools/research into the risk assessment process.

Charge Question 4: *To what extent do the ORD intramural projects, the extramural Science To Achieve Results (STAR) centers, and the five stated CTRP management priorities described in the FY09-12 implementation plan combine to efficiently support the goal of providing high-throughput decision support tools for screening and assessing chemical exposure, hazard, and risk to human health?*

Ultimately, the implementation of the CTRP's goals is going to require an iterative approach to developing models as the data available grows in both quantity and complexity. To do this, the CTRP must develop methods that can be used to make verifiable predictions and to generate appropriate data to test those predictions. This clearly is going to require developing a strong partnership that reaches beyond the boundaries of the CTRP and takes advantage of other existing programs, including ORD intramural projects, other intramural EPA projects, the extramural STAR Centers, and partnerships with other agencies.

The FY09-12 implementation plan lays out these interactions and the role that they will play in helping to direct and develop the CTRP. This document serves as a useful reference and guide as to the roles and activities of all parties contributing to the CTRP. The new start programs outlined for the ORD laboratories and centers are well leveraged with the CTRP. Results from those projects have the potential to be quickly incorporated into ongoing CTRP activities, providing for effective and efficient use of R&D efforts. Continuing cooperation with the STAR Centers and the innovations that are taking place at these university sites will enhance the

development and robustness of the decision support tools under development. In total, NCCT interactions with other ORD laboratories and the STAR Centers are essential as they will provide not only the starting data needed to fully develop the CTRP, but also some sampling of the potential users of the systems that will be essential for providing feedback and ideas for the next iteration of tools.

In its implementation plan, the CTRP lays out five priority management areas:

1. Toxicity Predictions and Chemical Prioritizations Incorporating Exposure
2. Strengthening Cross-ORD Collaborations
3. Tox21: A Federal Partnership Transforming Toxicology
4. Communicating Computational Toxicology
    a. EPA Program Office Training and Implementation of Computational Tools
    b. Communities of Practice for Chemical Prioritization and Exposure Science
5. Developing Clients for Virtual Tissues

Overall~~n the whole~~, it is evident that addressing the five priority management areas will further support program efficiency and effectiveness, and help sustain the progress the CTRP has achieved to date. The C~~Subc~~ommittee members agree that the combined programs outlined in the implementation plan are key contributions that are needed to maintain the CTRP on a path to achieve the stated goals.

Commented [J3]: I thought we agreed that these are not subcommittees but rather just BOSC committees

The following paragraphs present some of the BOSC's observations about some challenges along with suggestions for meeting them that are relevant to the Program's continued progress.

The CTRP needs to be more integrative, both internally and externally, to ensure all parties are working from common assumptions, data development schedules, and deliverable planning. As an example, at the review, there were instances where the theoretical and modeling groups noted that they needed access to quantitative data to inform certain aspects of their models while in the same session the experimental groups were presenting precisely those data. Some of this speaks to the stage of the various projects, but as the ultimate goal is to have computational tools that are useful for informing risk assessment, these groups need to begin to work together more closely.

It will be important to detail specific roles for the STAR Centers as part of the integrated approach to managing the Program's mission. One of the omissions in this management plan is the role that the STAR Centers will play in the future. Although the STAR Centers and their importance in the CTRP's future plans are laid out elsewhere in the implementation plan, it will be important to include these into a more integrated approach to managing the CTRP's mission. T~~T~~he c~~Subc~~ommittee's understanding is that this omission is deliberate and related to the fact that the STAR Program funding is independent of the CTRP budget; consequently, CTRP management cannot rely on it. The Centers, however, play a key role in the activities of the CTRP; therefore, it is crucial to have a firm plan for their continued contribution. The CTRP should consider the implications for the research program if a STAR Center is not renewed, and prepare some contingencies on how the research gaps might be filled if these key contributors

are lost from the research program.

As the overall CTRP continues to progress, there should be a higher priority on incorporation of ecological receptors and greater focus on assessment of exposure factors. The ~~Subcommittee~~Committee noted an absence of ecological health as an endpoint for the high-throughput decision support tools for screening and assessing chemical exposure, hazard, and risk. To become fully integrated and supportive of the Agency's regulatory activities, the CTRP will have to move into the field of ecological risk assessment at some point. Acknowledging this need and developing a forward plan to incorporate it as part of the CTRP should be part of the longer term plan.

The ExpoCast initiative, as described in the Hubal & Egeghy poster, is an important direction for the NCCT to pursue. Expansion of this effort to include real exposure and outcomes data, as well as the additional development of software resources to take advantage of these data for exposure and outcome predictions, should be a priority of the NCCT.

Training

One area that deserves particular comment in regard to the management plan is its training component, and in particular the training of postdoctoral fellows. The ~~Subcommittee~~Committee was universally impressed with the quality of the postdoctoral fellows and their work. The BOSC encourages the CTRP to continue its emphasis on training postdoctoral fellows because these scientists have the potential to be ambassadors to the rest of the community to help extend the understanding and acceptance of the types of computational tools the CTRP is trying to develop, and in doing so, ultimately help to improve those tools and their efficacy.

Quality Assurance for Software and Models

As the CTRP continues and expands its efforts to develop complex models and software systems, such as the virtual tissues, testing and quality assurance of these tools becomes even more important. Software and model testing is the practice of probing for errors (or "bugs"), typically by using a structured set of manually constructed inputs to generate a list of specific performance errors. Static testing of software and models involves inspection of the source code and formulas, usually in a structured fashion called a walk-through. Dynamic testing involves executing the code or model on a set of test inputs. Structured input sets can test how these tools perform in the face of boundary conditions (e.g., null or very long inputs), and systematically vary combinations of representative inputs. These methods are called "black box" tests because they do not require any knowledge of the implementation. Dynamic methods that make use of knowledge of the implementation are called "white box" approaches; for example, code coverage metrics that test what proportion of the source code is reached while processing an input suite. "White box" methods can provide information about how to improve test suites themselves. Open source approaches to development of software can be exploited as an extension of "white box" methods bringing large communities of software engineers to the evaluation of code.

A frequent cause of software or model failures is a lack of compatibility with another

application, an operating system, or Web browser. Compatibility testing is particularly important in a plan for distributed development, where the separately developed components must be compatible with each other. Test suites that exercise all aspects of each component's interface to the others address this issue in distributed software and model development.

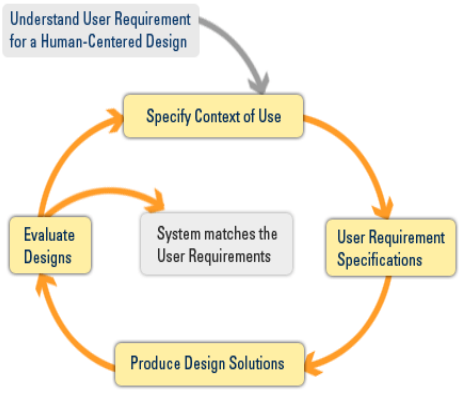
Testing in each of the proposed computational research areas is complicated by the fact that structured input sets and code coverage metrics have blind spots in these advanced systems. The range of possible inputs to a tissue model or natural language processing system cannot be exhausted by a structured test set. For these reasons, a software and model testing approach must be augmented with a sophisticated evaluation approach that probes how the tools and systems produced work in the hands of users.

User-Centered Design

Developing computer models and software systems to support complex and incompletely specified activities (such as modeling toxicity, assessing risk, and prioritizing exposures) is a difficult task. Users are generally unable to specify *a priori* what would be most useful for a computer system to do, although generally they do not have trouble describing what they like and do not like about any particular implementation. Furthermore, customer communities are heterogeneous, with needs that vary with analytical goals, methodological approaches, the types of data being analyzed, the amount and quality of relevant background knowledge available, and a wide variety of other factors. Customers and developers both can be frustrated by this seemingly circular need to produce models or software before defining requirements, which can themselves change in different circumstances and as new databases and software are produced.

User-centered design is an approach that grounds the process of design in information about the people who will use the product. There is an international standard (ISO 13407: Human-centered design process) that defines the general approach. Figure 1 shows the iterative nature of this process: starting with an understanding of the context of use (who will use it, under what conditions, to what ends) leads to a set of requirements that must be met, which in turn leads to a design solution, which then is evaluated through usability testing with actual users, which may lead to additional understanding of the context of use, and so on.

Figure 1: The ISO User-Centered Design Process



Qualitative methodology is a proven approach for effectively characterizing and explaining such issues as how scientists and policy makers make meaning while proceeding through complex analyses and how they inscribe visualized representations of knowledge into these problem-solving practices (Neressian, 2008). Field

observations can reveal the ways in which users generate preferred, new, or augmented analytical practices as they progressively use a new technology (Mirel, 2009; Vicente, 2002). Additionally, qualitative ethnographic methods can reveal ambiguities that scientists negotiate amid biological uncertainty and incomplete data and show the processes by which they disambiguate them. NSF recently issued a report stressing the importance of the study of scientists' research processes through qualitative methods (Lamont & White, 2005).

In general, trained observers using longitudinal fieldwork are better than domain experts themselves at discerning discrete steps in their own reasoning and the role that artifacts such as software play in facilitating their work (Schon, 1983; Hutchins, 1996). Field observations and semi-structured post hoc interviews can importantly complement focus groups or structured interviews or surveys for elucidating longitudinal views of scientists' and policy makers' thinking and behaving (Schon, 1983; Cresswell and Plano, 2006).

Recommendations:

- ✧ Be more integrative, both internally and externally, to ensure all parties are working from common assumptions, data development schedules, and deliverable planning.

- ✧ Expand outreach to the broader community, both within EPA and in the extramural community. This is not to say that the CTRP has not been effective in building a strong outreach program, but only that this needs to be a priority, and possibly a higher priority.

- ✧ Detail specific roles for the STAR Centers as part of the integrated approach to managing the Program's mission.

- ✧ Place a higher priority on incorporation of ecological receptors and greater focus on assessment of exposure factors.

- ✧ Develop a forward, longer term plan to incorporate the field of ecological risk assessment as part of the CTRP.

- ✧ Expand the ExpoCast program to include real exposure and outcomes data, as well as the additional development of software resources to take advantage of these data for exposure and outcome predictions. This should be a priority of the Center.

- ✧ Continue training postdoctoral fellows because these scientists have the potential to be ambassadors to the rest of the community to help extend the understanding and acceptance of the types of computational tools the CTRP is trying to develop, and in doing so, ultimately help to improve those tools and their efficacy.

- ✧ Highlight quality assurance for software and models with a specific testing approach augmented with a sophisticated evaluation approach that probes how the systems produced work in the hands of users.

- ✧ Promote "user-centered design", an approach that grounds the process of design in

information about the people who will use the product.

Charge Question 5: *The NCCT was established as an organization with a 5-year charter ending in February 2010, which would continue dependent on: 1) meeting established goals; and 2) having continuing mission-critical goals and objectives. What recommendation(s) can you provide the Agency regarding continuation of the NCCT as an established organization, and the criticality of its goals and objectives to EPA?*

The BOSC strongly supports action by EPA to make the NCCT permanent. It is clear from EPA's Strategic Plan for Evaluating the Toxicity of Chemicals that computational toxicology will be integral to the future of toxicology, risk assessment, and regulatory decision-making by the Agency. EPA will not be able to fulfill its strategy, or indeed its mission, without significant expertise and an active research program in computational toxicology. The NCCT has made significant contributions during the short time it has been in existence. The Center's work products have had an impact on Agency activities. Center products such as DSSTox, Actor, and ToxRefDB have been of great assistance to EPA program offices and to the toxicology community at large. The longer-term projects underway at the Center have been productive and have demonstrated their potential value to EPA. The staff of the NCCT has proved that the structure of the Center—a core of strong expertise with additional expertise leveraged through collaborations outside the NCCT—is an ideal organizational structure in a resource-scarce environment. The BOSC recommends in the strongest terms that the NCCT be made permanent.

The CTRP has been enormously successful in developing the tools and resources necessary to bring computational predictions to the science of toxicology and risk assessment. In the process, the staff members have learned many of the lessons needed to move forward, including many of the reasons why simply building the framework they have in place has been so challenging. In moving forward, the Program cannot rest on its laurels, but must continue to address these fundamental infrastructure problems and questions, building on what it already has done. The challenge will be to do this, which can become all-absorbing, while continuing to achieve its goals and leverage resources effectively.

The CTRP also has assembled a tremendous intellectual resource, ranging from the more senior personnel involved in the Program to the staff and postdoctoral trainees. It is this intellectual infrastructure as much as the data, databases, and software that represent the real value of the Program. The various scientists who are the backbone of the Program are extraordinarily focused on the end goals and while there is always room for improvement in any program, it is these scientists and their commitment that will assure that the CTRP retains its focus.

Commented [J4]: Is this correct as rewritten?

One recommendation would be the establishment of performance metrics. Although this group has been publishing at a reasonable rate, the primary goal is not academic publication but rather the development of tools and resources for informing risk assessment, and there are potential objectives that can be used to assess the relative impact of these tools on the field, including Web hits, software downloads, and the citation rate of these tools in publications and grants. Although some of this was presented at the review, the Program should establish a subset of these as benchmarks and provide some measure of the historical change in these metrics during

future funding periods.

The NCCT has done a very good job of outreach to the risk assessors at EPA, and to other stakeholders within and outside the Agency. The members of the C~~Subc~~ommittee members saw ample evidence that the Program is focused on developing models and tools that can be applied to the mission of the Agency. Given that computational toxicology is on the leading edge of science, and that risk assessment decisions are generally made using established science, it will be important for the NCCT to continue to broadcast its message and to engage in constructive dialog with the prospective recipients of its work products.

Center management mentioned hiring a communications person to help disseminate NCCT's plans and products. The ~~Subcommittee~~Committee endorses this intention.

The NCCT has hired several highly accomplished scientists under Title 42 authorization. These individuals are providing critical leadership in computational toxicology and have contributed materially to both short-term and long-term initiatives. - They also have provided instant credibility and stature to the CTRP, which has been critical in recruiting post-docs and in establishing EPA as a leader in computational toxicology. - The ~~Subcommittee~~Committee members strongly believe that the Title 42 positions have been crucial in establishing direction for the CTRP and in providing continuing scientific leadership in this complex and cutting-edge area of research.

Recommendations:

- Establish performance metrics that track the development of tools and resources for informing chemical prioritization, toxicity testing, and risk assessment.

- Before CTRP presents its views on various chemicals in the coming years, it was recommended that the CTRP randomly identify 10 percent of the predictions of the models for comparison to the results of traditional toxicological testing so as to insure that there is no bias by the researchers prior to claiming that the the models are accurate predictors of adverse effects.

- Continue to meet with customers, clients, and stakeholders on a regular basis to ensure that the Program is meeting the needs of the risk assessors and risk managers in the Agency.

Formatted: Indent: Left: 0.5", No bullets or numbering

Formatted: No bullets or numbering

In conclusion, the BOSC members believe that the NCCT is making exceptional progress towards achieving its mission and the Board is pleased to provide advice on this important Center. The BOSC looks forward to future opportunities to provide timely advice to guide and improve the NCCT and its programs.

Sincerely,

Gary S. Sayler, Ph.D.
Chair, BOSC

Cresswell JW and Plano-Clark VL. (2006). Designing and Conducting Mixed Methods of Research. Thousand Oaks, CA: SAGE Publications.

Hutchins E. (1996). Cognition in the wild. MIT Press.

Lamont M and White P. (2005). Workshop on Interdisciplinary Standards for Systematic Qualitative Research (http://www.nsf.gov/sbe/ses/soc/ISSQR_workshop_rpt.pdf).

Mirel B. (2001). Usability for dynamic work: interface strategies that make or break support for complex problem solving. In: Proceedings of 2001 Usability Professionals Association Conference, June 25-29, 2001, Lake Las Vegas, NV.

Nersessian N. (2008). Creating Scientific Concepts. MIT Press.

Schon DA. (1983). The Reflective Practitioner: How Professionals Think in Action. Basic Books.

Vicente K. (2002). Ecological interface design: progress and challenges. *Human Factors* 44:62-78.